

# LINKING THE CIRCLE AND THE SIEVE: RAMANUJAN- FOURIER SERIES

H. Gopalkrishna Gadiyar and R. Padma  
AU-KBC Research Centre, M.I.T. Campus of Anna University  
Chromepet, Chennai 600 044 India  
E-mail: {gadiyar, padma}@au-kbc.org

## Abstract

Currently the circle and the sieve methods are the key tools in analytic number theory. In this paper the unifying theme of the two methods is shown to be Ramanujan - Fourier series.

**Introduction.** The two well known methods in additive number theory are the circle method and the sieve method. The circle method is based on using a generating function (See Section 3) and noting along with Ramanujan and Hardy that the rational points on the circle contribute most and then through estimates showing that the contribution from the other points is small. The other method is the sieve method which looks entirely different and tries to use probabilistic arguments by roughly giving a density of occurrence to multiples of numbers. In this note we show that the Ramanujan - Fourier series unite two different streams of thought, the analytic and the probabilistic.

To explain this in an intuitive way, we consider the twin prime and the related Goldbach problem and break the paper into three parts: psychological, logical and chronological. The psychological part conveys the intuition behind the argument. The logical part displays the techniques needed. The chronological part gives a whiggish(!) history of the subject (as far as we could trace with available resources).

**1. Psychological.** Let us assume that the primes follow some distribution. Then the twin prime problem would be deducible by studying the autocorrelation function of that distribution. The argument would require that the distribution have a Fourier series.

An autocorrelation function of a process is defined to be the expectation of the product of two random variables obtained by observing the process at different times. The well-known Wiener - Khintchine formula states a relationship between two important characteristics of a random process: the power spectrum of the process and the correlation function of the process [12]. It is used practically to extract hidden periodicities in seemingly random phenomena. The twin prime problem asks whether there are infinitely many prime pairs of the form  $(p, p + h)$  and how these pairs are distributed. One immediately notes that this is a problem of finding autocorrelation of primes and hence there should be a corresponding Wiener - Khintchine formula .

**2. Logical.** We state here the conventional Wiener - Khintchine formula , explain Ramanujan - Fourier series of an arithmetical function and the Wiener - Khintchine formula for an arithmetical function which leads to the conjecture of Hardy and Littlewood on the twin prime problem. We also give the stochastic interpretation of the Riemann zeta-function due to Rota.

**2.1 The Wiener - Khintchine formula.** The Wiener - Khintchine formula [12] basically says that if

$$f(t) = \sum_n f_n e^{i\lambda_n t}, \quad (1)$$

then

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t + \tau) \overline{f(t)} dt = \sum_n |f_n|^2 e^{i\lambda_n \tau}. \quad (2)$$

The left hand side of (2) is called an autocorrelation function. The right hand side is nothing but the power spectrum.

**2.2 Ramanujan - Fourier series.** The Ramanujan - Fourier series of an arithmetical function  $a(n)$  is an expansion of the form

$$a(n) = \sum_{q=1}^{\infty} a_q c_q(n), \quad (3)$$

where

$$c_q(n) = \sum_{\substack{k=1 \\ (k,q)=1}}^q e^{2\pi i \frac{k}{q} n}, \quad (4)$$

and  $(k, q)$  denotes the greatest common divisor of  $k$  and  $q$ .  $c_q(n)$  is known as the Ramanujan sum and  $a_q$ 's are called the Ramanujan - Fourier coefficients. The Ramanujan - Fourier coefficient  $a_q$ 's are evaluated as follows. Denote by  $M(g)$  the mean value of an arithmetical function  $g$ , that is,

$$M(g) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} g(n). \quad (5)$$

For  $1 \leq k \leq q$ ,  $(k, q) = 1$ , let  $e_{\frac{k}{q}}(n) = e^{2\pi i \frac{k}{q} n}$ ,  $(n \in \mathcal{N})$ . If  $a(n)$  is an arithmetical function with expansion (3), then

$$a_q = \frac{1}{\phi(q)} M(a c_q) = \frac{1}{\phi(q)} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} a(n) c_q(n), \quad (6)$$

where  $\phi(q)$  denotes the Euler totient function. Also,

$$M(e_{\frac{k}{q}} \overline{e_{\frac{k'}{q'}}}) = \begin{cases} 1, & \text{if } \frac{k}{q} = \frac{k'}{q'}, \\ 0, & \text{if } \frac{k}{q} \neq \frac{k'}{q'}. \end{cases} \quad (7)$$

The Ramanujan sum  $c_q(n)$  satisfies some nice arithmetical properties which are stated below.

(i)  $c_q(n)$  is multiplicative. That is,

$$c_{qq'}(n) = c_q(n) c_{q'}(n) \text{ if } (q, q') = 1. \quad (8)$$

(ii) If  $p$  is prime, then

$$c_p(n) = \begin{cases} -1, & \text{if } p \nmid n, \\ p-1, & \text{if } p|n, \end{cases} \quad (9)$$

where  $a|b$  means  $a$  divides  $b$  and  $a \nmid b$  means  $a$  does not divide  $b$ .

In number theory, it is customary to use the von Mangoldt function  $\Lambda(n)$  in the place of characteristic function of primes where  $\Lambda(n)$  is defined as follows.

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n = p^k, p, \text{ a prime } k, \text{ a positive integer} \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

The Ramanujan - Fourier expansion for  $\frac{\phi(n)}{n}\Lambda(n)$  can be obtained using the properties of  $c_q(n)$  and it is given by

$$\frac{\phi(n)}{n}\Lambda(n) = \sum_{q=1}^{\infty} \frac{\mu(q)}{\phi(q)} c_q(n), \quad (11)$$

where  $\mu(q)$  is the Möbius function.

Also, for a given integer  $h$ ,

$$C(h) \stackrel{\text{def}}{=} \sum_{q=1}^{\infty} \frac{\mu^2(q)}{\phi^2(q)} c_q(h) = \begin{cases} 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{\substack{p|h \\ p>2}} \left(\frac{p-1}{p-2}\right), & \text{if } h \text{ is even,} \\ 0, & \text{if } h \text{ is odd,} \end{cases} \quad (12)$$

where  $\prod_p$  denotes the product over primes.

The proof of (12) follows from the multiplicative properties of  $\mu(q)$ ,  $\phi(q)$  and  $c_q$ . For a given  $h$ , the series on the left hand side of (12) is absolutely convergent and hence has the Euler product expansion

$$\prod_p \left(1 + \frac{\mu^2(p)}{\phi^2(p)} c_p(h)\right) = \prod_{p \nmid h} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{p|h} \left(1 + \frac{1}{p-1}\right), \quad (13)$$

where we have used the property (ii) of  $c_q$  given by (9). When  $h$  is odd, the infinite product on the right hand side of (13) is 0 and when  $h$  is even, it is equal to

$$2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{\substack{p|h \\ p>2}} \left(1 + \frac{1}{p-1}\right) \left(1 - \frac{1}{(p-1)^2}\right)^{-1}, \quad (14)$$

which on simplification gives the right hand side of (12).

### 2.3 The Wiener - Khintchine formula for an arithmetical function.

The Wiener - Khintchine formula for an arithmetical function  $a(n)$  having the Ramanujan - Fourier series (3) can be stated as follows.

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} a(n)a(n+h) = \sum_{q=1}^{\infty} a_q^2 c_q(h). \quad (15)$$

(15) holds for those arithmetical functions whose Ramanujan - Fourier series are absolutely and uniformly convergent. But the Ramanujan - Fourier series for  $\frac{\phi(n)}{n} \Lambda(n)$  does not belong to this class of functions. However, if we assume the truth of the theorem in the most general case, it would give the formula conjectured by Hardy and Littlewood [9] which we now state.

*There are infinitely many prime pairs  $p, p+h$  for every even integer  $h$  and if  $\pi_h(N)$  denotes the number of prime pairs less than  $N$ , then*

$$\pi_h(N) \sim C(h) \frac{N}{\log^2 N}. \quad (16)$$

where  $C(h)$  is given by (12).

The Wiener - Khintchine formula for  $\frac{\phi(n)}{n} \Lambda(n)$  is given by

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \frac{\phi(n)}{n} \Lambda(n) \frac{\phi(n+h)}{n+h} \Lambda(n+h) = \sum_{q=1}^{\infty} \frac{\mu^2(q)}{\phi^2(q)} c_q(h) \quad (17)$$

$$= C(h). \quad (18)$$

Let

$$\Psi(h, N) = \sum_{n \leq N} \frac{\phi(n)}{n} \Lambda(n) \frac{\phi(n+h)}{n+h} \Lambda(n+h). \quad (19)$$

Then (17) implies that

$$\frac{\Psi(h, N)}{N} \sim C(h). \quad (20)$$

The terms of the sum on the left hand side of (17) are non zero if and only if both  $\Lambda(n)$  and  $\Lambda(n + h)$  are prime powers, say  $p^k$  and  $q^l$  respectively, for primes  $p$  and  $q$ . Passing from (20) to (16) is a standard exercise in number theory. See [6].

We now give the strong numerical evidence based on computerized calculations which indicate the plausibility of the Wiener - Khintchine formula [6]. For  $h = 2, 4$  and  $6$ , the formula (20) gives

$$\begin{aligned} \Psi(2, N) &\sim C(2) N \\ \Psi(4, N) &\sim C(4) N \\ \Psi(6, N) &\sim C(6) N \end{aligned}$$

Since  $C(4) = C(2)$  and  $C(6) = 2C(2)$ , there should be approximately equal numbers of prime power pairs differing by 2 and by 4, but about twice as many differing by 6. We have given below the actual values of  $\Psi(h, N)$  and also the ratio  $C(h)/\left(\frac{\Psi(h, N)}{N}\right)$  (fourth column) for  $h = 2, 4$  and  $6$  and

$N$  upto  $10^6$ . We have used the value of  $C(2) = 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \sim 1.320323632$  to compute the ratio.

**Table 1**

N	$\Psi(2, N)$	$\frac{\Psi(2, N)}{N}$	Ratio
100000	131522.552204	1.315226	1.003876
200000	264287.347531	1.321437	0.999158
300000	393317.025988	1.311057	1.007068
400000	525523.270611	1.313808	1.004959
500000	654557.716460	1.309115	1.008562
600000	789035.163302	1.315059	1.004004
700000	919941.157912	1.314202	1.004658
800000	1049182.174335	1.311478	1.006745
900000	1180813.946552	1.312015	1.006332
1000000	1312843.985016	1.312844	1.005697

**Table 2**

N	$\Psi(4, N)$	$\frac{\Psi(4, N)}{N}$	Ratio
100000	130212.335085	1.302123	1.013977
200000	260492.247225	1.302461	1.013714
300000	390320.617781	1.301069	1.014799
400000	527155.226011	1.317888	1.001848
500000	653649.051733	1.307298	1.009964
600000	789177.513123	1.315296	1.003823
700000	923982.224287	1.319975	1.000264
800000	1054670.388142	1.318338	1.001506
900000	1180133.117590	1.311259	1.006913
1000000	1307978.775955	1.307979	1.009438

**Table 3**

N	$\Psi(6, N)$	$\frac{\Psi(6, N)}{N}$	Ratio
100000	261289.742091	2.612897	1.010620
200000	523391.109218	2.616956	1.009053
300000	787393.641752	2.624645	1.006097
400000	1056087.319082	2.640218	1.000162
500000	1316336.875799	2.632674	1.003029
600000	1579274.310330	2.632124	1.003238
700000	1839327.388416	2.627611	1.004961
800000	2104826.034045	2.631033	1.003654
900000	2368450.398104	2.631612	1.003434
1000000	2631198.406265	2.631198	1.003591

**2.4 Stochastic interpretation of Rota.** In this section we point out the fact that Rota's stochastic interpretation of the Riemann - zeta function [19], [1] is based on profinite characters. We will see that Rota considers the group  $C_\infty$  of rational numbers modulo 1 and crucially bases his arguments (given below) on  $C_\infty^*$ , the group of characters of  $C_\infty$ . *Both Ramanujan and Rota are using the same tool as  $e^{2\pi i \frac{k}{q} n}$  of Ramanujan is a concrete realization of the characters of the profinite group of Rota.* It seems that Rota did not notice the Ramanujan connection. As Sieve methods [7] are based on probabilistic

considerations, the circle and the sieve methods are linked through the ideas of Rota and Ramanujan.

Consider  $A$  a subset of positive integers  $\mathcal{N}$ , then the arithmetic density is defined as

$$\text{dens}(A) = \lim_{n \rightarrow \infty} \frac{1}{n} \|A \cap \{1, 2, \dots, n\}\|, \quad (21)$$

whenever the limit exists. It is immediately obvious that the  $\text{dens}(\mathcal{N})=1$  and  $\text{dens}(A_p) = \frac{1}{p}$  where  $A_p$  is the set of multiples of  $p$ . For technical reasons (lack of countable additivity) the right way to go about it is to choose a number  $s > 1$  define the measure of a positive integer  $n$  to be  $\frac{1}{n^s}$ . Then it turns out that the measure of  $\mathcal{N}$  is equal to  $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ . Hence a countably additive measure  $P_s$  on the set  $\mathcal{N}$  defined as

$$P_s(A) = \frac{1}{\zeta(s)} \sum_{n \in A} \frac{1}{n^s}. \quad (22)$$

Further the fundamental property

$$P_s(A_p \cap A_q) = P_s(A_p)P_s(A_q) = \frac{1}{p^s q^s} \quad (23)$$

can be checked and  $\lim_{s \rightarrow 1} P_s(A) = \text{dens}(A)$ . That is, the arithmetic density though not a probability is the limit of probabilities.

Rota then gives a combinatorial twist to the problem. He considers a cyclic group  $C_r$  of order  $r$ . Every character  $\chi$  of the group  $C_r$  has a kernel which is a subgroup of  $C_r$ . Further every sequence  $\chi_1, \dots, \chi_s$  of characters of  $C_r$  has a joint kernel which is a subgroup of  $C_r$ . By a joint kernel of a sequence of characters we mean the intersection of their kernels. Suppose one were to choose a sequence of  $s$  characters independently and randomly and ask the question: what is the probability that the joint kernel is a certain subgroup  $C_n$  of  $C_r$ ? It can be seen that with probability  $\frac{1}{n}$  the kernel of a randomly chosen character will contain the subgroup  $C_n$  as there are  $r$  characters of the group  $C_r$  and  $\frac{r}{n}$  such characters will vanish on  $C_n$ . So the probability of the joint kernel will contain  $C_n$  is  $\frac{1}{n^s}$ . Let  $P_{C_n}$  denote the probability that the joint kernel of characters shall be  $C_n$ . It follows that  $\frac{1}{n^s} = \sum_{n|d|r} P_{C_d}$  This



is based on the fact that the partially ordered set of subgroups of a cyclic group  $C_r$  and the partially ordered set of the divisors of  $r$  are isomorphic. Next using Möbius inversion and change of variable  $d = nj$  it follows that

$$P_{C_n} = \sum_{n|d|r} \mu\left(\frac{d}{n}\right) \frac{1}{d^s} \quad (24)$$

$$P_{C_n} = \frac{1}{n^s} \sum_j \mu(j) \frac{1}{j^s} \quad (25)$$

The variable  $j$  ranges over a subset of divisors of  $r$ . If the sum ranged over all positive integers  $j$  we would get the probabilistic interpretation

$$\frac{1}{n^s} \frac{1}{\zeta(s)}. \quad (26)$$

This is done by replacing the finite cyclic group  $C_n$  by a profinite cyclic group. Take the group  $C_\infty$  of rational numbers modulo 1. For every positive integer  $n$  the group  $C_\infty$  has unique finite subgroup  $C_n$  of order  $n$ . The character group  $C_\infty^*$  of  $C_\infty$  is a compact group, and the Haar measure is a probability measure. The group  $C_\infty^*$  is the desired profinite group. Mimicking the argument made earlier it can be seen that

$$\frac{1}{n^s} = \sum_{n|d} P_{C_d} \quad (27)$$

and by Möbius inversion

$$P_{C_n} = \sum_{n|d} \mu\left(\frac{d}{n}\right) \frac{1}{d^s} = \frac{1}{n^s} \frac{1}{\zeta(s)}. \quad (28)$$

The characters of  $C_\infty$  are given by  $e^{2\pi i \frac{k}{q} n}$  and so one immediately observes that the connection between  $C_\infty^*$  of Rota and the Ramanujan - Fourier expansion.

**3. Chronological.** In this section we will re-analyze the circle method and extract from it the essence which leads to great simplification and an intuitive understanding of the issues involved. We will also give the historical roots of probabilistic ideas in Hardy's work. The two streams will merge because of Rota's profinite characters being related to Ramanujan - Fourier series.

**3.1 Meaning of the circle method.** We begin by retracing what we feel should have been the train of thought of Hardy, Littlewood and Ramanujan, as far as it is possible, from their published researches. The first problem to be attacked by the circle method was the partition problem. As is well known, this reduces to understanding the generating function

$$\sum_{n=0}^{\infty} p(n) x^n = \frac{1}{(1-x)(1-x^2)(1-x^3)\dots}. \quad (29)$$

The key observation that led to the development of the circle method was that  $p(n)$  could be written as

$$\frac{1}{2\pi i} \int_C \frac{1}{(1-z)(1-z^2)(1-z^3)\dots} z^{-n-1} dz, \quad (30)$$

where  $C$  is the circle  $|z| = r$  and  $r < 1$ . Hardy and Ramanujan observed that the unit circle is covered by infinity of singularities corresponding to the poles of the generating function at all the rational points on the unit circle. This led them to carry out some extremely subtle and delicate analysis which gave remarkably accurate numerical agreement with available results. Having got this success, Hardy along with Littlewood went on to attack the other problems in additive number theory like the Waring's problem and the Goldbach problem. In each case they got approximate formulae in terms of what is called the singular series. At this point, Hardy and Littlewood seem to have shifted their focus to the complex analytic aspects of the circle method whereas Ramanujan wrote what is probably the one of the deepest papers in number theory in which he introduced the concept of what is now called the Ramanujan - Fourier expansion [18].

What Ramanujan does is to show by simple, yet ingenious methods that a wide range of arithmetical functions have Ramanujan - Fourier expansions, that is, an expansion of the form (3). He however does not indicate any formula for getting the Ramanujan - Fourier coefficients  $a_q$  which are the backbone of Fourier analysis. This was done by R. D. Carmichael [3] a little later. He showed that the Ramanujan - Fourier coefficients  $a_q$  can be got by the formula (6). M. Kac, E. R. Van Kampen and A. Wintner [10], [11] have pointed out the almost periodic nature of the Ramanujan-Fourier expansions. See also [13] and [20].

In the mean time, Hardy and Littlewood realised that their singular series could be evaluated because it was actually a Ramanujan - Fourier expansion. To do this, Hardy used the techniques he had developed in one of his most important papers [8] in which he stressed the multiplicative nature of  $c_q(n)$  ((8) and (9)) which enabled him to get closed form formula for the singular series. Using the properties of  $c_q$  Hardy also obtained a Ramanujan - Fourier expansion of  $\frac{\phi(n)}{n}\Lambda(n)$  given by (11).

As historical aside, we would like to remark that clues of Ramanujan - Fourier expansion playing an important role in additive number theoretic problems are available in the work of Glaisher [9] in which a simple minded partial fraction expansion yields this.

$$\frac{1}{(1-x)(1-x^2)(1-x^3)} = \frac{1}{6(1-x)^3} + \frac{1}{4(1-x)^2} \quad (31)$$

$$+ \frac{17}{72(1-x)} + \frac{1}{8(1+x)} \quad (32)$$

$$+ \frac{1}{9(1-\omega x)} + \frac{1}{9(1-\omega^2 x)}, \quad (33)$$

where  $\omega$  and  $\omega^2$  denote the two complex cube roots of unity. If

$$\frac{1}{(1-x)(1-x^2)(1-x^3)} = 1 + \sum_{n=1}^{\infty} r(n)x^n, \quad (34)$$

then

$$r(n) = \frac{(n+3)^2}{12} - \frac{7}{72} + \frac{(-1)^n}{8} + \frac{2}{9} \cos\left(\frac{2n\pi}{3}\right). \quad (35)$$

Over the last half century, several refinements have been made in attacking additive number theory problems. In the circle method, as applied by Hardy and Littlewood to the Goldbach problem, strong hypotheses regarding primes in arithmetic progressions have to be made. However, ultimately what is got is a main term in the formula which is given in terms of the Ramanujan - Fourier series and an error term. The hard part is to control the error term. At this point, we would like to indicate what we consider the line that Ramanujan would have taken had he not lost interest in the Ramanujan - Fourier expansion.

H. Wilf [21] remarks that “A generating function is a clothesline on which we hang up a sequence of numbers for display”. But is a clothesline necessary, that is, do we need a generating function at all? Rather than considering the generating function and then applying the circle method which yields a Ramanujan - Fourier expansion, can we not directly get the answer? What should be done is a clean shave with Occam’s Razor.

Consider a typical additive number theory problem. We would begin with a subset  $A$  of positive integers and let  $a(n)$  denote its characteristic function. If the problem is to find out  $r(n)$  which is the number of ways  $n$  can be written as a sum of  $k$  elements of the set  $A$ , we would extract the  $n^{th}$  coefficient of  $(\sum a(n)x^n)^k$ . This corresponds to taking the  $k$ -fold convolution. In other words,

$$r(n) = \sum_{i_1+i_2+\dots+i_k=n} a(i_1)a(i_2)\dots a(i_k).$$

All the analysis carried out over the past few years finally leads to an approximate formula for this convolution.

If  $a(n)$  has the Ramanujan - Fourier expansion (3), then

$$r(n) \sim \binom{n+k-1}{k-1} \sum_{q=1}^{\infty} a_q^k c_q(n). \quad (36)$$

This formula is immediately understandable as the usual relationship between the convolution and multiplication which exists for Fourier series. The combinatorial factor which appears is due to the fact that for almost periodic functions [2] only approximate formulae exist and this combinatorial factor is necessary there.

The reason why the rational points dominate is simply because of the fact that the  $a(n)$  have Ramanujan - Fourier expansions which give rise to the simple poles at all the rational points on the unit circle. For,

$$\sum_{n=1}^{\infty} a(n)x^n = \sum_{n=1}^{\infty} \sum_{q=1}^{\infty} \sum_{\substack{k=1 \\ (k,q)=1}}^q a_q e^{2\pi i \frac{k}{q}n} x^n \quad (37)$$

$$= \sum_{q=1}^{\infty} \sum_{\substack{k=1 \\ (k,q)=1}}^q a_q \sum_{n=1}^{\infty} \left( e^{2\pi i \frac{k}{q}n} x \right)^n \quad (38)$$

$$= \sum_{q=1}^{\infty} \sum_{\substack{k=1 \\ (k,q)=1}}^q a_q \frac{e^{2\pi i \frac{k}{q} x}}{1 - e^{2\pi i \frac{k}{q} x}}. \quad (39)$$

For example, for the Goldbach problem, it is known that  $\frac{\phi(n)\Lambda(n)}{n}$  has a Ramanujan - Fourier expansion (11). It is this  $\frac{\mu(q)}{\phi(q)}$  which is got by using the properties of primes in arithmetic progressions in the circle method. Hence if an approximate formula for convolution can be proved rigorously, the Goldbach problem can be solved. In other methods there are many sources of error and the methods do not tell us why things work. Further they involve many hypotheses. Hence a return to the ideas of Ramanujan seems to be absolutely necessary to clarify and simplify the circle method.

Let us now look at the Goldbach conjecture more closely. If  $r(n)$  denotes the number of ways  $n = 2m$  can be expressed as a sum of two primes, then from (11) and (36),

$$r(n) \sim C(n) \frac{n}{\log^2 n}, \quad (40)$$

where  $C(n)$  is given by (12). This formula was conjectured by Hardy and Littlewood. But Sylvester was the first mathematician to suggest an asymptotic formula for  $r(n)$ . He conjectured that

$$r(n) \sim 2\pi(n) \prod \left( \frac{p-2}{p-1} \right), \quad (41)$$

where  $\pi(n)$  denotes the number of primes up to  $n$  and the product extends over all odd primes  $p \leq n$  and  $p \nmid n$ . There is no evidence in the literature as to how he arrived at this formula. We know by prime number theorem that

$$\pi(n) \sim \frac{n}{\log n}, \quad (42)$$

and by Merten's formula

$$\prod_{p \leq x} \left( 1 - \frac{1}{p} \right) \sim \frac{e^{-\gamma}}{\log x}, \quad (43)$$

where  $\gamma$  is Euler's constant. This shows that

$$\prod_{p \leq \sqrt{n}} \left( \frac{p-2}{p-1} \right) = \prod_{p \leq \sqrt{n}} \left( 1 - \frac{1}{(p-1)^2} \right) \prod_{p \leq \sqrt{n}} \left( 1 - \frac{1}{p} \right) \sim \frac{2Ae^{-\gamma}}{\log n}, \quad (44)$$

where

$$A = \prod_{p > 2} \left( 1 - \frac{1}{(p-1)^2} \right), \quad (45)$$

where the product runs over all odd primes. Thus

$$r(n) \sim \frac{4Ae^{-\gamma}n}{\log^2 n} \prod_{\substack{p|n \\ p > 2}} \left( \frac{p-2}{p-1} \right). \quad (46)$$

One may call this as 'Sylvester's formula'. This formula is clearly wrong. It is right in its most important terms, but the constant  $4Ae^{-\gamma}$  is not correct and the correct formula should just have  $2A$  as the constant factor. However, the sieve argument used by Merlin and Brun gives an asymptotic formula similar to the one in given by (46) which is given below.

**3.2 Probabilistic Interpretation of the Goldbach conjecture.** Here one forms the table

$$\begin{array}{cccccccccccc} 1, & 2, & 3, & 4, & 5, & 6, & 7, & \dots, & n-1 & (m) \\ n-1, & n-2, & n-3, & n-4, & n-5, & n-6, & n-7, & \dots, & 1 & (m') \end{array}$$

where  $n$  is an even number and the table read vertically gives the decompositions  $n = m + m'$  of  $n$  into positive integers  $m$  and  $m'$ . Let us now perform the sieve of Eratosthenes on both rows of the table sieving with the first  $l$  primes starting from the right of the lower row. A decomposition  $m + m'$  is considered erased when either  $m$  or  $m'$  is erased.

Two cases occur here. If the prime  $p|n$ , then the erasures in the second row fall immediately below the corresponding erasures in the first, and the corresponding erasures in the first, and the number of decompositions erased is approximately  $\frac{n}{p}$ . If  $p \nmid n$ , then the erasures never correspond, and now the number of decompositions is approximately  $\frac{2n}{p}$ . If  $r_l(n)$  corresponds to the number of decompositions  $n = m + m'$  such that  $m$  or  $m'$  are not divisible by any of the first  $l$  primes, we have

$$r_l(n) = n \prod_{p|n, p \leq p_l} \left( 1 - \frac{1}{p} \right) \prod_{p \nmid n, p \leq p_l} \left( 1 - \frac{2}{p} \right). \quad (47)$$

This formula is correct for a fixed  $l$ . But if we assume the truth of this formula for  $p_l \sim \sqrt{n}$ , then we have  $r(n) = r_l(n)$  and

$$r(n) = \frac{1}{2}n \prod_{p|n} \left( \frac{p-1}{p-2} \right) \prod_{p \leq \sqrt{n}} \left( 1 - \frac{2}{p} \right), \quad (48)$$

where  $p = 2$  is now excluded. By Merten's formula, if  $3 \leq p \leq \sqrt{n}$ ,

$$\prod \left( 1 - \frac{2}{p} \right) = \prod \left( 1 - \frac{1}{(p-1)^2} \right) \left( 1 - \frac{1}{p} \right)^2 \sim \frac{16Ae^{-2\gamma}}{\log^2 n}. \quad (49)$$

Thus,

$$r(n) \sim \frac{8Ae^{-2\gamma}n}{\log^2 n} \prod_{p|n} \left( \frac{p-1}{p-2} \right). \quad (50)$$

This is the formula to which Brun's argument naturally leads. Like Sylvester's formula this is also wrong but correct up to the factor of  $4e^{-2\gamma}$ .

A similar probabilistic argument for the twin prime conjecture has been lucidly explained by Polya [17].

Let us now recall the probabilistic interpretation of  $\zeta(s)$  given in Section 2.3 and note that the  $e^{2\pi i \frac{k}{q}n}$  of Ramanujan is a concrete realization of the characters of the profinite group of Rota. Thus the ideas of circle method and sieve method can be linked through the work of Rota via Ramanujan - Fourier expansion.

**3.3 Related ideas.** Physicists have carried out statistical analysis of primes recently and shown that the distribution of primes is related to  $1/f$  noise. Marek Wolf [22] has shown that the power spectrum displays the  $1/f^\beta$  behaviour with the exponent  $\beta \sim 1.64$  by performing the Fourier transform of the "signal" given by the number of primes contained in the successive intervals of equal length  $l = 2^{16} = 65536$  up to  $N = 2^{38} \sim 2.74910^{11}$ . Here this slope  $\beta$  does not depend on the length of the sampled intervals, which suggests some kind of self-similarity in the distribution of primes. From Hardy and Ramanujan we know that primes have Ramanujan - Fourier expansion. Michel Planat [16] has used Ramanujan - Fourier series to signal processing (which is unrelated to number theory.)

Before we conclude we make two comments of historical nature.

1. In a paper on time-series analysis [5], A. Einstein has discussed the autocorrelation function and its relationship to the spectral content of a time series many years before Wiener and Khintchine.
2. It is well known that Hardy had true antipathy to probability theory. In [4] Diaconis shows that despite his antipathy, Hardy contributed significantly to modern probability. “His work with Ramanujan begat probabilistic number theory. His work on Tauberian theorems and divergent series has probabilistic proofs and interpretations. Finally Hardy spaces are a central ingredient in stochastic calculus.” He refers to Hardy’s work with Littlewood on the Goldbach and prime  $k$ -tuples conjecture as an instance of his insightful probabilistic thinking. “Hardy’s papers give a sophisticated development of conjectured asymptotics which suggest that at least one of the authors was quite familiar with probabilistic heuristics. .... Hardy refers to what I would call probabilistic reasoning as a priori judgment of common sense in his expository account (Hardy (1922, page 2).”

Thus the two streams of thought of Hardy one, the heuristic probabilistic (sieve) versus the other, the rigorous analytic (circle) are merged into one stream via Ramanujan - Fourier series.

**Conclusion: Ontological and epistemological.** The reader might have noticed that though there is no proof in the strict mathematical sense of the word, there is convincing calculation which shows the plausibility of the conjecture. The philosophy of this paper is what Roddam Narasimha [15] calls computational positivism. Here we have a model for the primes  $s$  behaving like random variables and based on this model a numerical calculation is carried out which agrees with “experimental” data. This is the typical method used repeatedly in physics where one guesses a model and carries out a numerical calculation.

Both the circle method and the sieve method lead to Ramanujan - Fourier series. In the circle method, poles occurring at rational points lead to Ramanujan - Fourier series. In the sieve method, the profinite group characters give the right density on the integers. These two methods are therefore linked through Ramanujan - Fourier series.

**Dedication.** It is a pleasure to write an article in honour of Professor K.



Ramachandra on his 70<sup>th</sup> birthday. He belongs to the pre-independence nationalist school of scientists like Ramanujan (whom he admires.) He has been a symbol of resistance to neo-colonial fashions and an advocate of indigenous styles of research.

## References

1. K. S. Alexander, K. Baclawski, G.C. Rota, *A stochastic interpretation of the Riemann zeta-function*, Proc. Natl. Acad. Sci. USA, **90**, (1993) 697-99.
2. H. Bohr, *Almost periodic functions*, Chelsea Publishing Company, New York, 1951.
3. R. D. Carmichael, Expansions of arithmetical functions in infinite series, Proc. London Math. Soc. (2) 34 (1932), 1-26.
4. P. Diaconis, *G.H. Hardy and Probability* Bull London Math. Soc. **34**, (2002) 385-402.
5. A. Einstein, *Method for the determination of the statistical values of observations concerning quantities subject to irregular fluctuations*, Archives des Sciences et Naturelles, **37**, (1914) 254-256.
6. H. G. Gadiyar and R. Padma, *Ramanujan-Fourier series, the Wiener-Khintchine formula and the distribution of prime pairs*, Physica **A** 269(1999) 503-510.
7. G. Greaves, *Sieves in Number Theory*, Springer-Verlag, Berlin-Heidelberg, 2001.
8. G. H. Hardy, Note on Ramanujan's trigonometrical function  $c_q(n)$  and certain series of arithmetical functions, Proc. Camb. Phil. Soc. **20**, (1921) 263 - 271.
9. G. H. Hardy and J. E. Littlewood, 'Some problems of Partition Numerorum'; III: On the expression of a number as a sum of primes, Acta Math. **44**, (1922) 1-70.
10. M. Kac, *Almost periodicity and the representation of integers as sums of squares*, Am. J. Math. **62**, (1940) 122 - 126.

11. M. Kac, E. R. Van Kampen and A. Wintner, *Ramanujan sums and almost periodic functions*, Am. J. Math. **62**, (1940) 107 - 114.
12. C. Kittel, *Elementary Statistical Physics*, John Wiley & Sons Inc., New York, 1958.
13. J. Knopfmacher, *Abstract Analytic Number Theory*, North-Holland/American Elsevier, 1975.
14. J. Kubilius, *Probabilistic Methods in the Theory of Numbers*, Vol. 11, Translations of Mathematical Monographs, American Mathematical Society Providence, Rhode Island, 1964.
15. Roddam Narasimha, *Axiomatism and computational positivism - Two mathematical cultures in pursuit of exact sciences*, Economic and Political Weekly, August 30, 2003, 3650-3656.
16. M. Planat, H. Rosu and S. Perrine, *Ramanujan sum for signal processing of low frequency noise* Phys. Rev. E **66** (2002) 56128.
17. G. Polya, *The Random Walks of George Polya*, Appendix 12, Heuristic Reasoning in the Theory of Numbers, 265-275.
18. S. Ramanujan, On certain trigonometrical sums and their applications in the theory of numbers, Trans. Camb. Phil. Soc. 22 (1918), 259 - 276.
19. G.C. Rota, *Combinatorial snapshots*, The Mathematical Intelligencer, **21** (1999) No.2, 8-14.
20. W. Schwarz and J. Spilker, *Mean values of Ramanujan expansions of almost even arithmetical functions*, In: Proc. 1974, Colloq. on Number Theory, Debrecen.
21. H. Wilf, *Generatingfunctionology*, Academic Press, 1990.
22. M. Wolf,  $1/f$  noise in the distribution of prime numbers, Physica A **241** (1997), 493-499.